EAST SEARCH

2/26/2007

L#	Hits	Search String	Databases
S16	3	S6 and (highest near2 gain)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S19	3	S6 and (select$3 with (stage or phase))	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S9	7	S6 and (candidate near2 feature)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S18	17	S6 and ((adjust$3 or modify$3) with model$3)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S17	39	S6 and (model$3 with feature)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S12	3	S6 and (rank$3 with feature)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S11	4	S6 and (rank$3 with gain)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S2	76	S1 and ("maximum entropy" near2 model$3)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S20	1	S6 and (gain with "upper bound")	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S1	10467	(language near2 model$3) or ("natural language" near2 (processing or model))	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S13	10	S6 and (order$3 with feature)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S15	2	S6 and ("top-ranked" with feature)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S7	23	S6 and (select$3 near2 feature)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S4	76	S3 and ("maximum entropy" near2 model$3)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S14	2	S6 and ("top ranked" with feature)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S3	10474	(language near2 model$3) or ("natural language" near2 (processing or model$3))	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S10	6	S6 and ((comput$3 or determin$3) with gain)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S6	93	S4 or S5	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S8	6	S6 and (gain with feature)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S24	1	S6 and (re-comput$3 with gain)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S25	6	S6 and ((conditional near2 probabilit$3) with feature)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S21	0	S6 and (gain with "upper limit")	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S22	0	S6 and (reevaluat$3 with gain)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S23	0	S6 and (re-evaluat$3 with gain)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S28	1	S6 and ("next-ranked" with feature)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S29	0	S6 and (gain with (pre-determined or pre-specified))	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S26	1	S6 and (gain with "uniform distribution")	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S27	9	S6 and ("uniform distribution")	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S5	82	S3 and (entropy near2 model$3)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S32	0	S6 and (re-us$3 with feature with gain)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S33	0	S6 and (reus$3 with feature with gain)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S30	1	S6 and (gain with (predetermined or prespecified))	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S31	1	S6 and ("top-ranked" with feature with number)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S34	0	S6 and (reus$3 with gain)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S35	8	S6 and (reusing or reused)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S36	59	S7 or S8 or S9 or S10 or S11 or S12 or S13 or S14 or S15 or S16 or S17 or S18 or S19 or S2	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S37	10474	(language near2 model$3) or ("natural language" near2 (processing or model$3))	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S38	76	S37 and ("maximum entropy" near2 model$3)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S39	82	S37 and (entropy near2 model$3)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S40	93	S38 or S39	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S41	23	S40 and (select$3 near2 feature)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB

S42	6	S40 and (gain with feature)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S43	7	S40 and (candidate near2 feature)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S54	1	S40 and (gain with "upper bound")	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S44	6	S40 and ((comput$3 or determin$3) with gain)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S45	4	S40 and (rank$3 with gain)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S46	3	S40 and (rank$3 with feature)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S47	10	S40 and (order$3 with feature)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S48	2	S40 and ("top ranked" with feature)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S49	2	S40 and ("top-ranked" with feature)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S50	3	S40 and (highest near2 gain)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S51	39	S40 and (model$3 with feature)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S52	17	S40 and ((adjust$3 or modify$3) with model$3)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S53	3	S40 and (select$3 with (stage or phase))	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S56	6	S40 and ((conditional near2 probabilit$3) with feature)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S58	9	S40 and ("uniform distribution")	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S55	1	S40 and (re-comput$3 with gain)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S59	1	S40 and ("next-ranked" with feature)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S57	1	S40 and (gain with "uniform distribution")	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S61	1	S40 and ("top-ranked" with feature with number)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S60	1	S40 and (gain with (predetermined or prespecified))	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S62	8	S40 and (reusing or reused)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S63	59	S41 or S42 or S43 or S44 or S45 or S46 or S47 or S48 or S49 or S50 or S51 or S52 or S53 c	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S82	1	S67 and (re-comput$3 with gain)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S79	17	S67 and ((adjust$3 or modify$3) with model$3)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S80	3	S67 and (select$3 with (stage or phase))	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S68	23	S67 and (select$3 near2 feature)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S77	3	S67 and (highest near2 gain)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S81	1	S67 and (gain with "upper bound")	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S83	6	S67 and ((conditional near2 probabilit$3) with feature)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S84	1	S67 and (gain with "uniform distribution")	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S86	1	S67 and ("next-ranked" with feature)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S85	9	S67 and ("uniform distribution")	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S88	1	S67 and ("top-ranked" with feature with number)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S89	8	S67 and (reusing or reused)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S87	1	S67 and (gain with (predetermined or prespecified))	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S90	59	S68 or S69 or S70 or S71 or S72 or S73 or S74 or S75 or S76 or S77 or S78 or S79 or S80 c	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S91	2	6,304,841.pn.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S92	1	S91 and ((remov$3 or eliminat$3 or discard$3) with feature)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S74	10	S67 and (order$3 with feature)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S64	10498	(language near2 model$3) or ("natural language" near2 (processing or model$3))	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S70	7	S67 and (candidate near2 feature)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S76	2	S67 and ("top-ranked" with feature)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S78	39	S67 and (model$3 with feature)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S71	6	S67 and ((comput$3 or determin$3) with gain)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S67	93	S65 or S66	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S66	82	S64 and (entropy near2 model$3)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S69	6	S67 and (gain with feature)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S65	76	S64 and ("maximum entropy" near2 model$3)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB

S72	4	S67 and (rank$3 with gain)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S75	2	S67 and ("top ranked" with feature)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S73	3	S67 and (rank$3 with feature)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S93	11701	(language near2 model$3) or ("natural language" near2 (processing or model))	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S94	89	S93 and ("maximum entropy" near2 model$3)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S96	4	S94 and (rank$3 with feature)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S97	4	S94 and (rank$3 with gain)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S98	6	S94 and ((comput$3 or determin$3 or calculat$3) with gain)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S99	7	S95 or S96 or S97 or S98	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S95	2	S94 and ("top-ranked" with feature)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB
S100	2	S94 and ("selective gain")	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB

10/613366 **Fuliang Weng et al.**

EAST SEARCH

2/26/2007

Results of search set S91:

Document Kind Codes	Title	Issue Date	Current OR	Abstract
US 20060178869 A1	Classification filter for processing data for creating a language model	20060810	704/10	
US 20060159507 A1	One-row keyboard	20060720	400/472	
US 20060123448 A1	Programming guide content collection and recommendation system for viewing on a portable	20060608	725/51	
US 20060123000 A1	Machine learning system for extracting structured records from web pages and other text sou	20060608	707/5	
US 20060095250 A1	Parser for natural language processing	20060504	704/9	
US 20060088356 A1	One-row keyboard and approximate typing	20060427	400/472	
US 20060074670 A1	Method and system for interactive conversational dialogue for cognitively overloaded device u	20060406	704/257	
US 20060074630 A1	Conditional maximum likelihood estimation of naive bayes probability models	20060406	704/9	
US 20060020448 A1	Method and apparatus for capitalizing text using maximum entropy	20060126	704/10	
US 20060018541 A1	Adaptation of exponential models	20060126	382/181	
US 20060015320 A1	Selection and use of nonstatistical translation components in a statistical machine translation	20060119	704/2	
US 20050283363 A1	Interactive manual, system and method for vehicles and other complex equipment	20051222	704/257	
US 20050256685 A1	Exponential priors for maximum entropy models	20051117	703/2	
US 20050256680 A1	Exponential priors for maximum entropy models	20051117	702/181	
US 20050237227 A1	Mention-synchronous entity tracking system and method for chaining mentions	20051027	341/1	
US 20050228643 A1	Discovery of parallel text portions in comparable collections of corpora and training using com	20051013	704/9	
US 20050171783 A1	Message recognition using shared language model	20050804	704/276	
US 20050165580 A1	Exponential priors for maximum entropy models	20050728	702/181	
US 20050055209 A1	Semantic language modeling and confidence measurement	20050310	704/255	
US 20050049852 A1	Adaptive and scalable method for resolving natural language ambiguities	20050303	704/9	
US 20050021317 A1	Fast feature selection method and system for maximum entropy modeling	20050127	703/2	
US 20050015251 A1	High-order entropy error functions for neural classifiers	20050120	704/232	
US 20040193401 A1	Linguistically informed statistical models of constituent structure for ordering in sentence reali	20040930	704/9	
US 20040064438 A1	Method for data and text mining and literature-based discovery	20040401	707/1	
US 20030236662 A1	Sequential conditional generalized iterative scaling	20031225	704/224	
US 20030126102 A1	Probabilistic record linkage model derived from training data	20030703	706/21	
US 20030074183 A1	Method and system for encoding and accessing linguistic frequency data	20030417	704/1	

US 20030055655 A1	Text processing system	20030320	704/276
US 20020188421 A1	Method and apparatus for maximum entropy modeling, and method and apparatus for natura	20021212	702/181
US 20020165716 A1	Error corrective mechanisms for consensus decoding of speech	20021107	704/255
US 20020111793 A1	Adaptation of statistical parsers based on mathematical transform	20020815	704/10
US 20020111780 A1	Probability model selection using information-theoretic optimization criterion	20020815	703/2
US 20020038207 A1	Systems and methods for word prediction and speech recognition	20020328	704/9
US 20020032549 A1	Determining and using acoustic confusability, acoustic perplexity and synthetic acoustic word	20020314	703/2
US 20010056344 A1	COMMAND BOUNDARY IDENTIFIER FOR CONVERSATIONAL NATURAL LANGUAGE	20011227	704/235
US 20010028744 A1	Method for processing nodes in 3D scene and apparatus thereof	20011011	382/232
US 7031910 B2	Method and system for encoding and accessing linguistic frequency data	20060418	704/10
US 7028038 B1	Method for generating training data for medical text abbreviation and acronym normalization	20060411	707/100
US 7010486 B2	Speech recognition system, training arrangement and method of calculating iteration values f	20060307	704/255
US 6961685 B2	Probability model selection using information-theoretic optimization criterion	20051101	703/2
US 6904405 B2	Message recognition using shared language model	20050607	704/235
US 6898320 B2	Method for processing nodes in 3D scene and apparatus thereof	20050524	382/232
US 6886010 B2	Method for data and text mining and literature-based discovery	20050426	707/3
US 6859774 B2	Error corrective mechanisms for consensus decoding of speech	20050222	704/255
US 6697769 B1	Method and apparatus for fast machine training	20040224	703/2
US 6640207 B2	Method and configuration for forming classes for a language model based on linguistic classe	20031028	704/9
US 6523019 B1	Probabilistic record linkage model derived from training data	20030218	706/45
US 6453292 B2	Command boundary identifier for conversational natural language	20020917	704/235
US 6415248 B1	Method for building linguistic models from a corpus	20020702	704/1
US 6304841 B1	Automatic construction of conditional exponential models from elementary features	20011016	704/2
US 6167377 A	Speech recognition language models	20001226	704/240
US 6107935 A	Systems and methods for access filtering employing relaxed recognition constraints	20000822	340/5.52
US 6049767 A	Method for estimation of feature gain and training starting point for maximum entropy/minimu	20000411	704/240
US 5991710 A	Statistical translation system with features based on phrases or groups of words	19991123	704/2
US 5839106 A	Large-vocabulary speech recognition using an integrated syntactic and semantic statistical la	19981117	704/257
US 5680511 A	Systems and methods for word recognition	19971021	704/257
JP 2002373163 A	METHOD AND APPARATUS FOR CREATING MAXIMUM ENTROPY MODEL AND METHO	20021226	
WO 2005008365 A	High quality feature selection method for maximum entropy modeling involves selecting top-r	20050127	
US 20020188421 A	Optimum entropy modeling method used for language processor in speech dialogue system,	20021212	

Terms used **language processing maximum entropy feature gain rank**

Found **90** of **198,146**

Sort results by: publication date
Display results: expanded form
Save results to a Binder
Search Tips
Open results in a new window

Try an Advanced Search
Try this search in The ACM Guide

Results 41 - 60 of 90 Result page: previous 1 2 **3** 4 5 next

Relevance scale

41 Why nitpicking works: evidence for Occam's Razor in error correctors

Dekai Wu, Grace Ngai, Marine Carpuat

August 2004 **Proceedings of the 20th international conference on Computational Linguistics COLING '04**

Publisher: Association for Computational Linguistics

Full text available: pdf(283.34 KB) Additional Information: full citation, abstract, references

Empirical experience and observations have shown us when powerful and highly tunable classifiers such as maximum entropy classifiers, boosting and SVMs are applied to language processing tasks, it is possible to achieve high accuracies, but eventually their performances all tend to plateau out at around the same point. To further improve performance, various error correction mechanisms have been developed, but in practice, most of them cannot be relied on to predictably improve performance on un ...

42 Verb phrase ellipsis detection using automatically parsed text

Leif Arda Nielsen

August 2004 **Proceedings of the 20th international conference on Computational Linguistics COLING '04**

Publisher: Association for Computational Linguistics

Full text available: pdf(170.05 KB) Additional Information: full citation, abstract, references

This paper describes a Verb Phrase Ellipsis (VPE) detection system, built for robustness, accuracy and domain independence. The system is corpus-based, and uses a variety of machine learning techniques on free text that has been automatically parsed using two different parsers. Tested on a mixed corpus comprising a range of genres, the system achieves a 72% F1-score. It is designed as the first stage of a complete VPE resolution system that is input free text, detects VPEs, and proceeds to find ...

43 Linguistically informed statistical models of constituent structure for ordering in sentence realization

Eric Ringger, Michael Gamon, Robert C. Moore, David Rojas, Martine Smets, Simon Corston-Oliver

August 2004 **Proceedings of the 20th international conference on Computational Linguistics COLING '04**

Publisher: Association for Computational Linguistics

Full text available: pdf(136.95 KB) Additional Information: full citation, abstract, references

We present several statistical models of syntactic constituent order for sentence

realization. We compare several models, including simple joint models inspired by existing statistical parsing models, and several novel conditional models. The conditional models leverage a large set of linguistic features without manual feature selection. We apply and evaluate the models in sentence realization for French and German and find that a particular conditional model outperforms all others. We employ a ...

44 High-performance tagging on medical texts

Udo Hahn, Joachim Wermter

August 2004 **Proceedings of the 20th international conference on Computational Linguistics COLING '04**

Publisher: Association for Computational Linguistics

Full text available: pdf(96.19 KB) Additional Information: full citation, abstract, references

We ran both Brill's rule-based tagger and TNT, a statistical tagger, with a default German newspaper-language model on a medical text corpus. Supplied with limited lexicon resources, TNT outperforms the Brill tagger with state-of-the-art performance figures (close to 97% accuracy). We then trained TNT on a large annotated medical text corpus, with a slightly extended tagset that captures certain medical language particularities, and achieved 98% tagging accuracy. Hence, statistical off-the-shelf ...

45 Acquiring the meaning of discourse markers

Ben Hutchinson

July 2004 **Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics ACL '04**

Publisher: Association for Computational Linguistics

Full text available: pdf(95.33 KB) Additional Information: full citation, abstract, references

This paper applies machine learning techniques to acquiring aspects of the meaning of discourse markers. Three subtasks of acquiring the meaning of a discourse marker are considered: learning its **polarity, veridicality**, and **type** (i.e. causal, temporal or additive). Accuracy of over 90% is achieved for all three tasks, well above the baselines.

46 Incremental parsing with the perceptron algorithm

Michael Collins, Brian Roark

July 2004 **Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics ACL '04**

Publisher: Association for Computational Linguistics

Full text available: pdf(141.53 KB) Additional Information: full citation, abstract, references

This paper describes an incremental parsing approach where parameters are estimated using a variant of the perceptron algorithm. A beam-search algorithm is used during both training and decoding phases of the method. The perceptron approach was implemented with the same feature set as that of an existing generative model (Roark, 2001a), and experimental results show that it gives competitive performance to the generative model on parsing the Penn treebank. We demonstrate that training a perceptr ...

47 Discriminative language modeling with conditional random fields and the perceptron algorithm

Brian Roark, Murat Saraclar, Michael Collins, Mark Johnson

July 2004 **Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics ACL '04**

Publisher: Association for Computational Linguistics

Full text available: pdf(219.59 KB) Additional Information: full citation, abstract, references

This paper describes discriminative language modeling for a large vocabulary speech recognition task. We contrast two parameter estimation methods: the perceptron algorithm, and a method based on conditional random fields (CRFs). The models are

encoded as deterministic weighted finite state automata, and are applied by intersecting the automata with word-lattices that are the output from a baseline recognizer. The perceptron algorithm has the benefit of automatically selecting a relatively small ...

48 A maximum entropy approach to species distribution modeling

Steven J. Phillips, Miroslav Dudík, Robert E. Schapire

July 2004 **Proceedings of the twenty-first international conference on Machine learning ICML '04**

Publisher: ACM Press

Full text available: pdf(163.78 KB) Additional Information: full citation, abstract, references, citings

We study the problem of modeling species geographic distributions, a critical problem in conservation biology. We propose the use of maximum-entropy techniques for this problem, specifically, sequential-update algorithms that can handle a very large number of features. We describe experiments comparing maxent with a standard distribution-modeling tool, called GARP, on a dataset containing observation data for North American breeding birds. We also study how well maxent performs as a function of ...

49 Head-Driven Statistical Models for Natural Language Parsing

Michael Collins

December 2003 **Computational Linguistics**, Volume 29 Issue 4

Publisher: MIT Press

Full text available: pdf(633.30 KB) Additional Information: full citation, abstract, references, citings, index terms

This article describes three statistical models for natural language parsing. The models extend methods from probabilistic context-free grammars to lexicalized grammars, leading to approaches in which a parse tree is represented as the sequence of decisions corresponding to a head-centered, top-down derivation of the tree. Independence assumptions then lead to parameters that encode the X-bar schema, subcategorization, ordering of complements, placement of adjuncts, bigram lexical dependencies, ...

50 Knowledge management session 2: semantic web: Learning cross-document structural relationships using boosting

Zhu Zhang, Jahna Otterbacher, Dragomir Radev

November 2003 **Proceedings of the twelfth international conference on Information and knowledge management CIKM '03**

Publisher: ACM Press

Full text available: pdf(145.37 KB) Additional Information: full citation, abstract, references, index terms

Multi-document discoure analysis has emerged with the potential of improving various information retrieval applications. Based on the newly proposed Cross-document Structure Theory (CST), this paper describes an empirical study that uses boosting to classify CST relationships between sentence pairs extracted from topically related documents. We show that the binary classifier for determining existence of structural relationships significantly outperforms the baseline. We also achieve promising r ...

Keywords: boosting, classification, cross-document structure, discourse analysis

51 Music: Polyphonic music modeling with random fields

Victor Lavrenko, Jeremy Pickens

November 2003 **Proceedings of the eleventh ACM international conference on Multimedia MULTIMEDIA '03**

Publisher: ACM Press

Full text available: pdf(240.18 KB) Additional Information: full citation, abstract, references, citings, index

terms

Recent interest in the area of music information retrieval and related technologies is exploding. However, very few of the existing techniques take advantage of recent developments in statistical modeling. In this paper we discuss an application of Random Fields to the problem of creating accurate yet flexible statistical models of polyphonic music. With such models in hand, the challenges of developing effective searching, browsing and organization techniques for the growing bodies of music col ...

52 Embedding web-based statistical translation models in cross-language information retrieval

Wessel Kraaij, Jian-Yun Nie, Michel Simard

September 2003 **Computational Linguistics**, Volume 29 Issue 3

Publisher: MIT Press

Full text available: pdf(381.29 KB) Additional Information: full citation, abstract, references, citings, index terms

Although more and more language pairs are covered by machine translation (MT) services, there are still many pairs that lack translation resources. Cross-language information retrieval (CLIR) is an application that needs translation functionality of a relatively low level of sophistication, since current models for information retrieval (IR) are still based on a bag of words. The Web provides a vast resource for the automatic construction of parallel corpora that can be used to train statistical ...

53 Posters: Music modeling with random fields

Victor Lavrenko, Jeremy Pickens

July 2003 **Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval SIGIR '03**

Publisher: ACM Press

Full text available: pdf(98.04 KB) Additional Information: full citation, references, index terms

54 A fast algorithm for feature selection in conditional maximum entropy modeling

Yaqian Zhou, Lide Wu, Fuliang Weng, Hauke Schmidt

July 2003 **Proceedings of the 2003 conference on Empirical methods in natural language processing - Volume 10**

Publisher: Association for Computational Linguistics

Full text available: pdf(267.62 KB) Additional Information: full citation, abstract, references, citings

This paper describes a fast algorithm that selects features for conditional maximum entropy modeling. Berger et al. (1996) presents an incremental feature selection (IFS) algorithm, which computes the approximate gains for all candidate features at each selection stage, and is very time-consuming for any problems with large feature spaces. In this new algorithm, instead, we only compute the approximate gains for the top-ranked features based on the models obtained from previous stages. Experimen ...

55 Single character Chinese named entity recognition

Xiaodan Zhu, Mu Li, Jianfeng Gao, Chang-Ning Huang

July 2003 **Proceedings of the second SIGHAN workshop on Chinese language processing - Volume 17**

Publisher: Association for Computational Linguistics

Full text available: pdf(176.35 KB) Additional Information: full citation, abstract, references, citings

Single character named entity (SCNE) is a name entity (NE) composed of one Chinese character, such as "[Abstract contained text which could not be captured.]" (*zhong1, China*) and "[Abstract contained text which could not be captured.]" (*e2, Russia*). SCNE is

very common in written Chinese text. However, due to the lack of in-depth research, SCNE is a major source of errors in named entity recognition (NER). This paper formulates the SCNE recognition within the source-channel model f ...

56 Identification of patients with congestive heart failure using a binary classifier: a case study

Serguei V. Pakhomov, James Buntrock, Christopher G. Chute

July 2003 **Proceedings of the ACL 2003 workshop on Natural language processing in biomedicine - Volume 13**

Publisher: Association for Computational Linguistics

Full text available: pdf(97.06 KB) Additional Information: full citation, abstract, references

This paper addresses a very specific problem that happens to be common in health science research. We present a machine learning based method for identifying patients diagnosed with congestive heart failure and other related conditions by automatically classifying clinical notes. This method relies on a Perceptron neural network classifier trained on comparable amounts of positive and negative samples of clinical notes previously categorized by human experts. The documents are represented as fea ...

57 Closing the gap: learning-based information extraction rivaling knowledge-engineering methods

Hai Leong Chieu, Hwee Tou Ng, Yoong Keok Lee

July 2003 **Proceedings of the 41st Annual Meeting on Association for Computational Linguistics - Volume 1 ACL '03**

Publisher: Association for Computational Linguistics

Full text available: pdf(115.36 KB) Additional Information: full citation, abstract, references, citings

In this paper, we present a learning approach to the scenario template task of information extraction, where information filling one template could come from multiple sentences. When tested on the MUC-4 task, our learning approach achieves accuracy competitive to the best of the MUC-4 systems, which were all built with manually engineered rules. Our analysis reveals that our use of full parsing and state-of-the-art learning algorithms have contributed to the good performance. To our knowledge, t ...

58 Confidence estimation for translation prediction

Simona Gandrabur, George Foster

May 2003 **Proceedings of the seventh conference on Natural language learning at HLT-NAACL 2003 - Volume 4**

Publisher: Association for Computational Linguistics

Full text available: pdf(342.14 KB) Additional Information: full citation, abstract, references

The purpose of this work is to investigate the use of machine learning approaches for confidence estimation within a statistical machine translation application. Specifically, we attempt to learn probabilities of correctness for various model predictions, based on the native probabilites (i.e. the probabilites given by the original model) and on features of the current context. Our experiments were conducted using three original translation models and two types of neural nets (single-layer and m ...

59 Inducing history representations for broad coverage statistical parsing

James Henderson

May 2003 **Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology - Volume 1 NAACL '03**

Publisher: Association for Computational Linguistics

Full text available: pdf(129.57 KB) Additional Information: full citation, abstract, references, citings

We present a neural network method for inducing representations of parse histories and

using these history representations to estimate the probabilities needed by a statistical left-corner parser. The resulting statistical parser achieves performance (89.1% F-measure) on the Penn Treebank which is only 0.6% below the best current parser for this task, despite using a smaller vocabulary size and less prior linguistic knowledge. Crucial to this success is the use of structurally determined soft bi ...

60 In question answering, two heads are better than one

Jennifer Chu-Carroll, Krzysztof Czuba, John Prager, Abraham Ittycheriah

May 2003 **Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology - Volume 1 NAACL '03**

Publisher: Association for Computational Linguistics

Full text available: pdf(102.68 KB) Additional Information: full citation, abstract, references, citings

Motivated by the success of ensemble methods in machine learning and other areas of natural language processing, we developed a multi-strategy and multi-source approach to question answering which is based on combining the results from different answering agents searching for answers in multiple corpora. The answering agents adopt fundamentally different strategies, one utilizing primarily knowledge-based mechanisms and the other adopting statistical techniques. We present our multi-level answer ...

Terms used **language processing maximum entropy feature gain rank**

Found **90** of **198,146**

Sort results by: publication date
Display results: expanded form
Save results to a Binder
Search Tips
Open results in a new window

Try an Advanced Search
Try this search in The ACM Guide

Results 61 - 80 of 90 Result page: previous 1 2 3 **4** 5 next

Relevance scale

61 Supervised and unsupervised PCFG adaptation to novel domains

Brian Roark, Michiel Bacchiani

May 2003 **Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology - Volume 1 NAACL '03**

Publisher: Association for Computational Linguistics

Full text available: pdf(139.23 KB) Additional Information: full citation, abstract, references

This paper investigates adapting a lexicalized probabilistic context-free grammar (PCFG) to a novel domain, using maximum *a posteriori* (MAP) estimation. The MAP framework is general enough to include some previous model adaptation approaches, such as corpus mixing in Gildea (2001), for example. Other approaches falling within this framework are more effective. In contrast to the results in Gildea (2001), we show F-measure parsing accuracy gains of as much as 2.5% for high accuracy lexical ...

62 Challenges in information retrieval and language modeling: report of a workshop held at the center for intelligent information retrieval, University of Massachusetts Amherst, September 2002

James Allan, Jay Aslam, Nicholas Belkin, Chris Buckley, Jamie Callan, Bruce Croft, Sue Dumais, Norbert Fuhr, Donna Harman, David J. Harper, Djoerd Hiemstra, Thomas Hofmann, Eduard Hovy, Wessel Kraaij, John Lafferty, Victor Lavrenko, David Lewis, Liz Liddy, R. Manmatha, Andrew McCallum, Jay Ponte, John Prager, Dragomir Radev, Philip Resnik, Stephen Robertson, Roni Rosenfeld, Salim Roukos, Mark Sanderson, Rich Schwartz, Amit Singhal, Alan Smeaton, Howard Turtle, Ellen Voorhees, Ralph Weischedel, Jinxi Xu, ChengXiang Zhai

April 2003 **ACM SIGIR Forum**, Volume 37 Issue 1

Publisher: ACM Press

Full text available: pdf(1.60 MB) Additional Information: full citation, citings, index terms, review

63 Special issue on special feature: Sufficient dimensionality reduction

Amir Globerson, Naftali Tishby

March 2003 **The Journal of Machine Learning Research**, Volume 3

Publisher: MIT Press

Full text available: pdf(266.18 KB) Additional Information: full citation, abstract, citings, index terms

Dimensionality reduction of empirical co-occurrence data is a fundamental problem in unsupervised learning. It is also a well studied problem in statistics known as the analysis of cross-classified data. One principled approach to this problem is to represent the data in low dimension with minimal loss of (mutual) information contained in the original data. In this paper we introduce an information theoretic nonlinear method for finding such a most informative dimension reduction. In contrast wi ...

64 Automatic labeling of semantic roles

Daniel Gildea, Daniel Jurafsky

September 2002 **Computational Linguistics**, Volume 28 Issue 3

Publisher: MIT Press

Full text available: pdf(573.51 KB) Additional Information: full citation, abstract, references, citings, index terms

*We present a system for identifying the semantic relationships, or **semantic roles**, filled by constituents of a sentence within a semantic frame. Given an input sentence and a target word and frame, the system labels constituents with either abstract semantic roles, such as* AGENT *or* PATIENT, *or more domain-specific semantic roles, such as* SPEAKER, MESSAGE, *and* TOPIC.*The system is based on statistical classifiers trained on roughly 50,000 sentences that were h ...*

65 The disambiguation of nominalizations

Maria Lapata

September 2002 **Computational Linguistics**, Volume 28 Issue 3

Publisher: MIT Press

Full text available: pdf(471.69 KB) Additional Information: full citation, abstract, references, citings, index terms

This article addresses the interpretation of nominalizations, a particular class of compound nouns whose head noun is derived from a verb and whose modifier is interpreted as an argument of this verb. Any attempt to automatically interpret nominalizations needs to take into account: (a) the selectional constraints imposed by the nominalized compound head, (b) the fact that the relation of the modifier and the head noun can be ambiguous, and (c) the fact that these constraints can be easily overr ...

66 Combining classifiers for Chinese word segmentation

Nianwen Xue, Susan P. Converse

September 2002 **Proceeding of the first SIGHAN workshop on Chinese language processing - Volume 18**

Publisher: Association for Computational Linguistics

Full text available: pdf(96.17 KB) Additional Information: full citation, abstract, references, citings

In this paper we report results of a supervised machine-learning approach to Chinese word segmentation. First, a maximum entropy tagger is trained on manually annotated data to automatically labels the characters with tags that indicate the position of character within a word. An error-driven transformation-based tagger is then trained to clean up the tagging inconsistencies of the first tagger. The tagged output is then converted into segmented text. The preliminary results show that this appro ...

67 Extracting the unextractable: a case study on verb-particles

Timothy Baldwin, Aline Villavicencio

August 2002 **proceeding of the 6th conference on Natural language learning - Volume 20 COLING-02**

Publisher: Association for Computational Linguistics

Full text available: pdf(173.70 KB) Additional Information: full citation, abstract, references, citings

This paper proposes a series of techniques for extracting English verb--particle

constructions from raw text corpora. We initially propose three basic methods, based on tagger output, chunker output and a chunk grammar, respectively, with the chunk grammar method optionally combining with an attachment resolution module to determine the syntactic structure of verb--preposition pairs in ambiguous constructs. We then combine the three methods together into a single classifier, and add in a number ...

68 Combining heterogeneous classifiers for word-sense disambiguation

Dan Klein, Kristina Toutanova, H. Tolga Ilhan, Sepandar D. Kamvar, Christopher D. Manning

July 2002 **Proceedings of the ACL-02 workshop on Word sense disambiguation: recent successes and future directions - Volume 8**

Publisher: Association for Computational Linguistics

Full text available: pdf(157.04 KB) Additional Information: full citation, abstract, references

This paper discusses ensembles of simple but heterogeneous classifiers for word-sense disambiguation, examining the Stanford-CS224N system entered in the SENSEVAL-2 English lexical sample task. First-order classifiers are combined by a second-order classifier, which variously uses majority voting, weighted voting, or a maximum entropy model. While individual first-order classifiers perform comparably to middle-scoring teams' systems, the combination achieves high performance. We discuss trade-of ...

69 Discriminative training methods for hidden Markov models: theory and experiments with perceptron algorithms

Michael Collins

July 2002 **Proceedings of the ACL-02 conference on Empirical methods in natural language processing - Volume 10 EMNLP '02**

Publisher: Association for Computational Linguistics

Full text available: pdf(295.08 KB) Additional Information: full citation, abstract, references, citings

We describe new algorithms for training tagging models, as an alternative to maximum-entropy models or conditional random fields (CRFs). The algorithms rely on Viterbi decoding of training examples, combined with simple additive updates. We describe theory justifying the algorithms through a modification of the proof of convergence of the perceptron algorithm for classification problems. We give experimental results on part-of-speech tagging and base noun phrase chunking, in both cases showing i ...

70 A hybrid approach to natural language web search

Jennifer Chu-Carroll, John Prager, Yael Ravin, Christian Cesar

July 2002 **Proceedings of the ACL-02 conference on Empirical methods in natural language processing - Volume 10 EMNLP '02**

Publisher: Association for Computational Linguistics

Full text available: pdf(118.21 KB) Additional Information: full citation, abstract, references, citings

We describe a hybrid approach to improving search performance by providing a natural language front end to a traditional keyword-based search engine. The key component of the system is iterative query formulation and retrieval, in which one or more queries are automatically formulated from the user's question, issued to the search engine, and the results accumulated to form the hit list. New queries are generated by relaxing previously-issued queries using transformation rules, applied in an ord ...

71 Machine learning in automated text categorization

Fabrizio Sebastiani

March 2002 **ACM Computing Surveys (CSUR)**, Volume 34 Issue 1

Publisher: ACM Press

Full text available: pdf(524.41 KB) Additional Information: full citation, abstract, references, citings, index terms

The automated categorization (or classification) of texts into predefined categories has

witnessed a booming interest in the last 10 years, due to the increased availability of documents in digital form and the ensuing need to organize them. In the research community the dominant approach to this problem is based on machine learning techniques: a general inductive process automatically builds a classifier by learning, from a set of preclassified documents, the characteristics of the categories. ...

Keywords: Machine learning, text categorization, text classification

72 Automatic verb classification based on statistical distributions of argument structure
Paola Merlo, Suzanne Stevenson
September 2001 **Computational Linguistics**, Volume 27 Issue 3
Publisher: MIT Press
Full text available: pdf(341.42 KB) Publisher Site Additional Information: full citation, abstract, references, citings

Automatic acquisition of lexical knowledge is critical to a wide range of natural language processing tasks. Especially important is knowledge about verbs, which are the primary source of relational information in a sentence---the predicate-argument structure that relates an action or state to its participants (i.e., who did what to whom). In this work, we report on supervised learning experiments to automatically classify three major types of English verbs, based on their argument structure--sp ...

73 XML-based data preparation for robust deep parsing
Claire Grover, Alex Lascarides
July 2001 **Proceedings of the 39th Annual Meeting on Association for Computational Linguistics ACL '01**
Publisher: Association for Computational Linguistics
Full text available: pdf(87.15 KB) Additional Information: full citation, abstract, references, citings

We describe the use of XML tokenisation, tagging and mark-up tools to prepare a corpus for parsing. Our techniques are generally applicable but here we focus on parsing Medline abstracts with the ANLT wide-coverage grammar. Hand-crafted grammars inevitably lack coverage but many coverage failures are due to inadequacies of their lexicons. We describe a method of gaining a degree of robustness by interfacing POS tag information with the existing lexicon. We also show that XML tools provide a soph ...

74 What is the minimal set of fragments that achieves maximal parse accuracy?
Rens Bod
July 2001 **Proceedings of the 39th Annual Meeting on Association for Computational Linguistics ACL '01**
Publisher: Association for Computational Linguistics
Full text available: pdf(69.87 KB) Additional Information: full citation, abstract, references, citings

We aim at finding the minimal set of fragments which achieves maximal parse accuracy in Data Oriented Parsing. Experiments with the Penn Wall Street Journal treebank show that counts of almost arbitrary fragments within parse trees are important, leading to improved parse accuracy over previous models tested on this treebank (a precision of 90.8% and a recall of 90.6%). We isolate some dependency relations which previous models neglect but which contribute to higher parse accuracy.

75 Beyond standard CFG parsing: New ranking algorithms for parsing and tagging: kernels over discrete structures, and the voted perceptron
Michael Collins, Nigel Duffy
July 2001 **Proceedings of the 40th Annual Meeting on Association for Computational Linguistics ACL '02**

Publisher: Association for Computational Linguistics
Full text available: pdf(137.02 KB) Additional Information: full citation, abstract, references, citings

This paper introduces new learning algorithms for natural language processing based on the perceptron algorithm. We show how the algorithms can be efficiently applied to exponential sized representations of parse trees, such as the "all subtrees" (DOP) representation described by (Bod 1998), or a representation tracking all sub-fragments of a tagged sentence. We give experimental results showing significant improvements on two tasks: parsing Wall Street Journal text, and named-entity extraction ...

76 The form is the substance: classification of genres in text
Nigel Dewdney, Carol VanEss-Dykema, Richard MacMillan
July 2001 **Proceedings of the workshop on Human Language Technology and Knowledge Management - Volume 2001**
Publisher: Association for Computational Linguistics
Full text available: pdf(64.60 KB) Additional Information: full citation, abstract, references

Categorization of text in IR has traditionally focused on topic. As use of the Internet and e-mail increases, categorization has become a key area of research as users demand methods of prioritizing documents. This work investigates text classification by format style, i.e. "genre", and demonstrates, by complementing topic classification, that it can significantly improve retrieval of information. The paper compares use of presentation features to word features, and the combination thereof, usin ...

77 Improving accuracy in word class tagging through the combination of machine learning systems
Hans van Halteren, Walter Daelemans, Jakub Zavrel
June 2001 **Computational Linguistics**, Volume 27 Issue 2
Publisher: MIT Press
Full text available: pdf(2.37 MB) Additional Information: full citation, abstract, references, citings
Publisher Site

We examine how differences in language models, learned by different data-driven systems performing the same NLP task, can be exploited to yield a higher accuracy than the best individual system. We do this by means of experiments involving the task of morphosyntactic word class tagging, on the basis of three different tagged corpora. Four well-known tagger generators (hidden Markov model, memory-based, transformation rules, and maximum entropy) are trained on the same corpus data. After comparis ...

78 Presentations: Two statistical parsing models applied to the Chinese Treebank
Daniel M. Bikel, David Chiang
October 2000 **Proceedings of the second workshop on Chinese language processing: held in conjunction with the 38th Annual Meeting of the Association for Computational Linguistics - Volume 12**
Publisher: Association for Computational Linguistics
Full text available: pdf(525.60 KB) Additional Information: full citation, abstract, references, citings

This paper presents the first-ever results of applying statistical parsing models to the newly-available Chinese Treebank. We have employed two models, one extracted and adapted from BBN's SIFT System (Miller et al., 1998) and a TAG-based parsing model, adapted from (Chiang, 2000). On sentences with ≤40 words, the former model performs at 69% precision, 75% recall, and the latter at 77% precision and 78% recall.

79 Comparison between tagged corpora for the named entity task
Chikashi Nobata, Nigel Collier, Jun'ichi Tsujii
October 2000 **Proceedings of the workshop on Comparing corpora - Volume 9**

Publisher: Association for Computational Linguistics
Full text available: pdf(700.37 KB) Additional Information: full citation, abstract, references, citings

We present two measures for comparing corpora based on information theory statistics such as gain ratio as well as simple term-class frequency counts. We tested the predictions made by these measures about corpus difficulty in two domains --- news and molecular biology --- using the result of two well-used paradigms for NE, decision trees and HMMs and found that gain ratio was the more reliable predictor.

80 Regular papers: Applying system combination to base noun phrase identification
Erik F. Tjong Kim Sang, Walter Daelemans, Hervé Déjean, Rob Koeling, Yuval Krymolowski, Vasin Punyakanok, Dan Roth
July 2000 **Proceedings of the 18th conference on Computational linguistics - Volume 2**
Publisher: Association for Computational Linguistics
Full text available: pdf(678.60 KB) Additional Information: full citation, abstract, references, citings

We use seven machine learning algorithms for one task: identifying base noun phrases. The results have been processed by different system combination methods and all of these outperformed the best individual result. We have applied the seven learners with the best combinator, a majority vote of the top five systems, to a standard data set and managed to improve the best published result for this data set.

Results 61 - 80 of 90 Result page: previous 1 2 3 **4** 5 next

Terms used **language processing maximum entropy feature gain rank**

Found **90** of **198,146**

Sort results by: publication date

Display results: expanded form

Save results to a Binder

Search Tips

Open results in a new window

Try an Advanced Search

Try this search in The ACM Guide

Results 81 - 90 of 90 Result page: previous 1 2 3 4 **5**

Relevance scale

81 The maximum entropy approach and probabilistic IR models

Warren R. Greiff, Jay M. Ponte

July 2000 **ACM Transactions on Information Systems (TOIS)**, Volume 18 Issue 3

Publisher: ACM Press

Full text available: pdf(246.45 KB) Additional Information: full citation, abstract, references, citings, index terms, review

This paper takes a fresh look at modeling approaches to information retrieval that have been the basis of much of the probabilistically motivated IR research over the last 20 years. We shall adopt a subjectivist Bayesian view of probabilities and argue that classical work on probabilistic retrieval is best understood from this perspective. The main focus of the paper will be the ranking formulas corresponding to the Binary Independence Model (BIM), presented originally by Roberston and Spar ...

Keywords: idf weighting, binary independence model, combination match, linked dependence, probability ranking principle

82 Web mining research: a survey

Raymond Kosala, Hendrik Blockeel

June 2000 **ACM SIGKDD Explorations Newsletter**, Volume 2 Issue 1

Publisher: ACM Press

Full text available: pdf(1.58 MB) Additional Information: full citation, citings, index terms

Keywords: Web, data mining, information extraction, information retrieval

83 Assigning function tags to parsed text

Don Blaheta, Eugene Charniak

April 2000 **Proceedings of the first conference on North American chapter of the Association for Computational Linguistics**

Publisher: Morgan Kaufmann Publishers Inc.

Full text available: pdf(602.13 KB) Additional Information: full citation, abstract, references, citings

It is generally recognized that the common nonterminal labels for syntactic constituents (NP, VP, etc.) do not exhaust the syntactic and semantic information one would like about

parts of a syntactic tree. For example, the Penn Tree-bank gives each constituent zero or more 'function tags' indicating semantic roles and other related information not easily encapsulated in the simple constituent labels. We present a statistical algorithm for assigning these function tags that, on text already parse ...

84 Information extraction: Information extraction research and applications: current progress and future directions

Andrew Kehler, Jerry R. Hobbs, Douglas Appelt, John Bear, Matthew Caywood, David Israel, Megumi Kameyama, David Martin, Claire Monteleoni

October 1998 **Proceedings of a workshop on held at Baltimore, Maryland: October 13-15, 1998**

Publisher: Association for Computational Linguistics

Full text available: pdf(1.24 MB) Additional Information: full citation, abstract, references

Analysts face a daunting task: they must accurately analyze, categorize, and assimilate a large body of information from a variety of sources and for a variety of domains of interest. The complexity of the task necessitates a variety of information access and extraction tools which technology up to this point has not been able to provide. SRI's TIPSTER Phase III project has focused on two major obstacles to the development of such tools: inadequate degrees of accuracy and portability. We begin b ...

85 Feature lattices for maximum entropy modelling

Andrei Mikheev

August 1998 **Proceedings of the 17th international conference on Computational linguistics - Volume 2 , Proceedings of the 36th annual meeting on Association for Computational Linguistics - Volume 2**

Publisher: Association for Computational Linguistics

Full text available: pdf(664.80 KB) Additional Information: full citation, abstract, references, citings

Maximum entropy framework proved to be expressive and powerful for the statistical language modelling, but it suffers from the computational expensiveness of the model building. The iterative scaling algorithm that is used for the parameter estimation is computationally expensive while the feature selection process might require to estimate parameters for many candidate features many times. In this paper we present a novel approach for building maximum entropy models. Our approach uses the featu ...

86 Memory-based learning: using similarity for smoothing

Jakub Zavrel, Walter Daelemans

July 1997 **Proceedings of the eighth conference on European chapter of the Association for Computational Linguistics , Proceedings of the 35th annual meeting on Association for Computational Linguistics**

Publisher: Association for Computational Linguistics

Full text available: pdf(764.37 KB) Publisher Site Additional Information: full citation, abstract, references, citings

This paper analyses the relation between the use of similarity in Memory-Based Learning and the notion of backed-off smoothing in statistical language modeling. We show that the two approaches are closely related, and we argue that feature weighting methods in the Memory-Based paradigm can offer the advantage of automatically specifying a suitable domain-specific hierarchy between most specific and most general conditioning information without the need for a large number of parameters. We report ...

87 Independence assumptions considered harmful

Alexander Franz

July 1997 **Proceedings of the eighth conference on European chapter of the Association for Computational Linguistics , Proceedings of the 35th annual**

meeting on Association for Computational Linguistics
Publisher: Association for Computational Linguistics
Full text available: pdf(695.22 KB) Publisher Site
Additional Information: full citation, abstract, references, citings

Many current approaches to statistical language modeling rely on independence assumptions between the different explanatory variables. This results in models which are computationally simple, but which only model the main effects of the explanatory variables on the response variable. This paper presents an argument in favor of a statistical approach that also models the interactions between the explanatory variables. The argument rests on empirical evidence from two series of experimetns concern ...

88 A new statistical parser based on bigram lexical dependencies
Michael John Collins
June 1996 **Proceedings of the 34th annual meeting on Association for Computational Linguistics**
Publisher: Association for Computational Linguistics
Full text available: pdf(737.31 KB) Publisher Site
Additional Information: full citation, abstract, references, citings

This paper describes a new statistical parser which is based on probabilities of dependencies between head-words in the parse tree. Standard bigram probability estimation techniques are extended to calculate probabilities of dependencies between pairs of words. Tests using Wall Street Journal data show that the method performs at least as well as SPATTER (Magerman 95; Jelinek et al. 94), which has the best published results for a statistical parser on this task. The simplicity of the approach me ...

89 A maximum entropy approach to natural language processing
Adam L. Berger, Vincent J. Della Pietra, Stephen A. Della Pietra
March 1996 **Computational Linguistics**, Volume 22 Issue 1
Publisher: MIT Press
Full text available: pdf(1.87 MB) Publisher Site
Additional Information: full citation, abstract, references, citings

The concept of maximum entropy can be traced back along multiple threads to Biblical times. Only recently, however, have computers become powerful enough to permit the widescale application of this concept to real world problems in statistical estimation and pattern recognition. In this paper, we describe a method for statistical modeling based on maximum entropy. We present a maximum-likelihood approach for automatically constructing maximum entropy models and describe how to implement this app ...

90 Parallel text search methods
Gerard Salton, Chris Buckley
February 1988 **Communications of the ACM**, Volume 31 Issue 2
Publisher: ACM Press
Full text available: pdf(1.53 MB)
Additional Information: full citation, abstract, references, citings, index terms, review

A comparison of recently proposed parallel text search methods to alternative available search strategies that use serial processing machines suggests parallel methods do not provide large-scale gains in either retrieval effectiveness or efficiency.

Nothing Found

Your search for **+"language processing" +"maximum entropy" +feature +gain +rank* +top-rank*** did not return any results.

You may want to try an Advanced Search for additional options.

Please review the Quick Tips below or for more information see the Search Tips.

Quick Tips

- Enter your search terms in lower case with a space between the terms.

  ```
  sales offices
  ```

  You can also enter a full question or concept in plain language.

  ```
  Where are the sales offices?
  ```

- Capitalize proper nouns to search for specific people, places, or products.

  ```
  John Colter, Netscape Navigator
  ```

- Enclose a phrase in double quotes to search for that exact phrase.

  ```
  "museum of natural history" "museum of modern art"
  ```

- Narrow your searches by using a + if a search term must appear on a page.

  ```
  museum +art
  ```

- Exclude pages by using a - if a search term must not appear on a page.

  ```
  museum -Paris
  ```

Combine these techniques to create a specific search query. The better your description of the information you want, the more relevant your results will be.

```
museum +"natural history" dinosaur -Chicago
```

Nothing Found

Your search for **+"language processing" +"maximum entropy" +feature +gain +top-rank*** did not return any results.

You may want to try an Advanced Search for additional options.

Please review the Quick Tips below or for more information see the Search Tips.

Quick Tips

- Enter your search terms in lower case with a space between the terms.

  ```
  sales offices
  ```

  You can also enter a full question or concept in plain language.

  ```
  Where are the sales offices?
  ```

- Capitalize proper nouns to search for specific people, places, or products.

  ```
  John Colter, Netscape Navigator
  ```

- Enclose a phrase in double quotes to search for that exact phrase.

  ```
  "museum of natural history" "museum of modern art"
  ```

- Narrow your searches by using a + if a search term must appear on a page.

  ```
  museum +art
  ```

- Exclude pages by using a - if a search term must not appear on a page.

  ```
  museum -Paris
  ```

Combine these techniques to create a specific search query. The better your description of the information you want, the more relevant your results will be.

```
museum +"natural history" dinosaur -Chicago
```

IEEE Xplore® RELEASE 2.2

Welcome United States Patent and Trademark Office

Results for "((('language modeling' and 'maximum entropy' and feature and gain and rank*)<in>metadata)" e-mail
Your search matched **0** documents.
A maximum of **100** results are displayed, **25** to a page, sorted by **Relevance** in **Descending** order.

» Search Options

View Session History

New Search

Modify Search

((('language modeling' and 'maximum entropy' and feature and gain and rank*)<in>me Search

Check to search only within this results set

Display Format: Citation Citation & Abstract

» Key

IEEE JNL	IEEE Journal or Magazine
IET JNL	IET Journal or Magazine
IEEE CNF	IEEE Conference Proceeding
IET CNF	IET Conference Proceeding
IEEE STD	IEEE Standard

No results were found.

Please edit your search criteria and try again. Refer to the Help pages if you need assistan search.

Indexed by Inspec®

Results for "(('language processing' and 'maximum entropy' and feature and gain and rank*)<in>metadata)" e-mail
Your search matched **0** documents.
A maximum of **100** results are displayed, **25** to a page, sorted by **Relevance** in **Descending** order.

» Search Options

View Session History

New Search

Modify Search

(('language processing' and 'maximum entropy' and feature and gain and rank*)<in>m Search

Check to search only within this results set

Display Format: Citation Citation & Abstract

» Key

IEEE JNL	IEEE Journal or Magazine
IET JNL	IET Journal or Magazine
IEEE CNF	IEEE Conference Proceeding
IET CNF	IET Conference Proceeding
IEEE STD	IEEE Standard

No results were found.

Please edit your search criteria and try again. Refer to the Help pages if you need assistan search.